An Analytical Model on the Blocking Probability of a Fault-Tolerant Network

Mathew P. Haynos and Yuanyuan Yang, Senior Member, IEEE

Abstract—The well-known Clos network has been extensively used for telephone switching, multiprocessor interconnection and data communications. Much work has been done to develop analytical models for understanding the routing blocking probability of the Clos network. However, none of the analytical models for estimating the blocking probability of this type of network have taken into account the very real possibility of the interstage links in the network failing. In this paper, we consider the routing between arbitrary network inputs and outputs in the Clos network in the presence of interstage link faults. In particular, we present an analytical model for the routing blocking probability of the Clos network which incorporates the probability of interstage link failure to allow for a more realistic and useful determination of the approximation of blocking probability. We also conduct extensive simulations to validate the model. Our analytical and simulation results demonstrate that for a relatively small interstage link failure probability, the blocking behavior of the Clos network is similar to that of a fault-free network, and indicate that the Clos network has a good fault-tolerant capability. The new integrated analytical model can guide network designers in the determination of the effects of network failure on the overall connecting capability of the network and allows for the examination of the relationship between network utilization and network failure.

Index Terms—Multi-stage interconnection networks, performance analysis, analytical model, fault tolerance, blocking probability, Clos network, random routing.

1 INTRODUCTION

Ongoing microprocessor developments have recently sparked interest in large-scale multiprocessors composed of hundreds or thousands of processors and in data communication networks allowing for delivery of advanced digital services. These developments have resulted in an increased focus on the connecting capabilities and the reliability of interconnection networks responsible for connecting the processing nodes in the network. To eliminate the need to support a direct connection from a given source node to each destination node, many interconnection networks are comprised of intermediate stages of switches used to route connection requests through. This type of interconnection network is generally referred to as multi-stage interconnection networks (MINs).

While there have been numerous designs proposed for multi-stage interconnection networks, each having its own merits, a network design proposed by Clos [1] originally for telephone networks continues to find applications today in multiprocessor interconnection networks and data communication networks. For example, the NEC ATOM switch for Broadband Integrated Services Digital Network (BISDN) is based on the three-stage Clos architecture [2] and, more recently, it was shown that the network in the IBM SP2 parallel computer is functionally equivalent to the Clos network [3].

It is not surprising that, as the number of components increases in a computing system (which an interconnection network is a significant part of), issues of fault tolerance and reliability as they relate to the interconnection network become ever more of a concern. Much of the study of fault-tolerant interconnection networks has been focused on network architectures which support some form of hardware redundancy, thereby allowing for limited types of network component failure. Examples include the Multi-path Omega network [4], Enhanced IADM network [5], d-dilated square banyan network [6], and Multibutterflies network [7]. Key distinguishing characteristics of fault-tolerant interconnection network architectures are the scope of component failure they allow for, with only a few encompassing a complete set, and in the number of faults tolerated [8].

Additionally, the reliability analysis of interconnection networks has been concentrated on computing some measure of overall network reliability and there have been many techniques proposed [9]. For example, [10], [11], [12] presented algorithms on the determination of fault-free path availability in several multi-path MINs. However, most models have not incorporated the various states an interstage link may be in. Of particular difficulty in the reliability analysis of interconnection networks is that many problems are computationally intractable.

Important work has also been done in establishing analytical models for understanding the connecting capability of an interconnection network; see, for example, [13], [14], [15], [16], [17], [18]. These models approximate the probability that an arbitrary connection request between a network input and a network output cannot be successfully routed through the network, i.e., the blocking probability, given various network component probabilities such as the probability that an input/output link is busy and the probability that an interstage link is busy. Few models,
The first stage in the three-stage network is called the input stage and consists of \( r \) input stage switches of size \( n \times m \). The second stage in the network is referred to as the middle stage and consists of \( m \) middle stage switches of size \( r \times r \). The third stage in the network is called the output stage and consists of \( r \) output stage switches of size \( m \times n \). Each input stage switch has exactly one connection to each of the \( m \) middle stage switches and this connection is referred to as an input-middle interstage link. Additionally, each middle stage switch has exactly one connection to each of the \( r \) output stage switches and this connection is referred to as a middle-output interstage link. An interstage link is said to be functional if it is capable of transmitting data and faulty otherwise. Furthermore, an interstage link is available if it is functional and not busy. The fault model we will use assumes that the interstage links in the network may fail and that these failures are permanent.

We consider the network capable of one-to-one or unicast communication. A connection request is a request to transmit data from an input port on an input stage switch to an output port on an output stage switch. A legal connection request is a connection request for which both the input port and the output port are not busy. Given a legal connection request, a routing algorithm attempts to route the connection in the network. A path is uniquely defined by an input port, an input stage switch, a middle stage switch, an output stage switch, and an output port.

A legal connection request is said to be satisfiable if a path can be found for which both the input-middle stage link and the middle-output stage link are available. Otherwise, the connection request is blocked. The blocking probability is the probability that a legal connection request is blocked. Finally, the network utilization is the percentage of the network input ports servicing active connections at any time.

In attempting to satisfy a legal connection request, a routing algorithm must be used to find a path through the network. The routing algorithm we will consider in this paper is a random-routing algorithm. A random-routing algorithm attempts to satisfy a legal connection request by randomly selecting an input-middle stage link from the set of all available input-middle stage links emanating from the input switch of the connection request. Next, the middle-output stage link from the chosen middle stage switch to the output stage switch of the connection request is checked to see if it is available. If it is, then a path is established in the network using the input-middle stage link and the middle-output stage link and the connection request is satisfied. If it is not, then another available input-middle stage link emanating from the input stage switch of the legal connection request is chosen and the process is repeated. If a path cannot be established in the network, then the connection request is blocked.

2 Preliminaries

The general Clos network is comprised of an input stage, an output stage, and an odd number of middle stages. Each stage consists of multiple switch modules. We will concentrate on the basic three-stage Clos network in this paper since any odd number stage networks with various switch sizes can be built in a recursive fashion from the three-stage networks. The schematic of a three-stage Clos network is depicted in Fig. 1. A switch module with \( n \) input ports and \( m \) output ports is referred to as an \( n \times m \) switch.
probabilistic results, which focus on analyzing the blocking probability of the network. Because the Clos network can potentially support many possible routings from a source node to a destination node, it is inherently more fault-tolerant of interstage link failure than many interconnection network designs. One may expect that the network has a good fault-tolerant capability in the presence of link faults. However, to our knowledge, no existing work has considered the fault-tolerant capability of the Clos network. In this paper, we address this issue in the context of probabilistic analysis. We will examine how interstage link failures affect the blocking probability of the network.

Much work has been done in the analysis of blocking probability for multistage interconnection networks. The work can be generally classified into two categories of interest. The first is from the viewpoint of reliability analysis and is referred to as the terminal reliability problem. Terminal reliability is defined as the probability that there is at least one operative path between a given pair of network input port and output port. For example, [10], [11], [12] presented algorithms for computing terminal reliability in several multipath MINs. The second category of work has been focused on establishing analytical models based on stochastic network parameters (i.e., network utilization) for the blocking probability of the network. Among the analytical models proposed for the Clos network in the literature that estimate blocking probability, two well-known and widely used models for random routing were proposed by Lee [14] and Jacobacue [13], respectively. Both models assume that the incoming traffic is uniformly distributed over the mi interstage links of the Clos network and the events that individual links are busy are independent. Lee gave the simplest method for analyzing the blocking probability of the Clos network. Let n be the number of ports on a particular input or output switch, m (\(\geq n\)) be the number of middle stage switches, a be the network utilization, \(p\) be the probability that an interstage link is busy and be defined as \(p = \frac{\alpha}{\beta}\), \(q\) be the probability that an interstage link is idle and be defined as \(q = 1 - p\). Then in Lee's model, the blocking probability, \(P_B\), is given by

\[
P_B = (1 - q^m)^n.
\]

A more accurate model was provided by Jacobsacue [13], and the blocking probability of the three-stage Clos network is calculated by the following formula:

\[
P_B = \frac{(2 - a)^{2m} - n^m}{n!(2n - m)!}.
\]

Both models, however, do not meet the deterministic nonblocking condition set forth by Clos [1] that, when \(m \geq 2n - 1\), the Clos network is nonblocking for arbitrary one-to-one communication. Recently, Yang [18] has presented a more accurate analytical model which still follows the same assumptions as in these two models but is proven to agree with the deterministic nonblocking condition as well. In this model, the probability that a connection request is not blocked for a three-stage Clos network is given by

\[
Pr\{\text{connection not blocked}\} = \sum_{n_1=0}^{n-1} \sum_{n_2=0}^{n-1} \min(n_1, n_2) \sum_{k=0}^{\max(0, n_1-n_2-m+1)} \left( \binom{n_1}{k} \binom{(n-k)-m}{n_2} \right) q^m \left( 1 - q^{m-n_1-n_2} \right) \left( \sum_{j=0}^{n-1} \binom{n}{j} p^j q^{n-j} \right)^2
\]

and the blocking probability is given by

\[
P_B = 1 - Pr\{\text{connection not blocked}\}.
\]

It is interesting to note that the categories seem somewhat discontinuous. Many models presented in the context of reliability analysis do not provide for a precise measure of blocking probability because of the view of network components as being either available or unavailable. The focus of these models has not been on incorporating network utilization parameters. Furthermore, analytical models presented for the blocking probability of the network incorporate network utilization, but do not support the notion of network component failure. We can expect then that analytical models which integrate the fact that network links can be either busy or faulty will provide a more realistic measure of network blocking probability.

4 A NEW PROBABILISTIC MODEL FOR INCORPORATING LINK FAILURE

In this section, we present an analytical model for the blocking probability of a three-stage Clos network which incorporates interstage link failure and allows for a more realistic measure of blocking probability. In general, determination of blocking probability in a multistage network is inherently complex and difficult. This is due to the fact that there are many possible paths to consider in a typical large network and the dependencies among links in the network lead to combinatorial explosion problems. Therefore, some approximations and assumptions are necessary for the calculations.

Let us first consider the network state depicted in Fig. 2, in which \(n_1\) input-middle interstage links from input switch \(i\) are busy, \(f_1\) input-middle interstage links from input switch \(i\) are faulty, \(n_2\) middle-output interstage links to output switch \(j\) are busy, and \(f_2\) middle-output interstage links to output switch \(j\) are faulty, where \(0 \leq n_1, n_2 \leq n - 1\), \(0 \leq f_1 \leq m - n_1\), \(f_2 \leq m - n_2\), and pairs of these interstage links are overlapped. A pair of interstage links is said to be overlapped if both links are either busy or faulty. Note that by definition \(n_1\) and \(f_1\) (and, similarly, \(n_2\) and \(f_2\)) are mutually exclusive, since an interstage link cannot be both busy and faulty. Further, it is important to note that the number of busy input-middle interstage links and the number of faulty input-middle interstage links are constrained by \(n_1\), that is, \(0 \leq n_1 + f_1 \leq m\). Similarly, for \(n_2\) and \(f_2\), we have \(0 \leq n_2 + f_2 \leq m\).

Having established the relationship between busy and faulty interstage links, it is evident that a new analytical model which incorporates interstage link failures can give a more realistic and useful measure of blocking probability of the Clos Network. In accounting for interstage link failures,
we need to define four events: $f_1$, the event that $f_1$ input-middle interstage links are faulty, $f_2$, the event that $f_2$ middle-output interstage links are faulty, $n_1$, the event that $n_1$ input-middle interstage links are busy and $n_2$, the event that $n_2$ middle-output interstage links are busy. Given these four events we can establish the probability that $k$ pairs of interstage links are overlapped. Looking at Fig. 2 shows that there are four states that a pair of input-middle interstage link and middle-output interstage link can be in: busy and busy, faulty and faulty, busy and faulty, and faulty and busy. Having defined overlapped links, we now determine the probability that some $k$ pairs of interstage links are overlapped, given events $n_1$, $f_1$, $n_2$, and $f_2$. We assume that the locations of faulty interstage links are independent and the faults are uniformly distributed over all interstage links. Under these assumptions and by taking advantage of the fact that an interstage link cannot both be busy and faulty, we have the following lemma concerning the overlapped interstage links.

**Lemma 1.** Given events $n_1$, $n_2$, $f_1$, and $f_2$, the probability that $k$ pairs of interstage links are overlapped is given by

$$
\Pr(k \text{ pairs of links overlapped} \mid n_1, n_2, f_1, f_2) = \frac{\binom{n_n + f_1}{k} \binom{n_m + f_2}{k} \binom{n_m - f_1}{n_n - f_1} \binom{n_m - f_2}{n_m - f_2}}{\binom{n_n + f_1 + f_2}{n_n + f_1 + f_2}}
$$

(5)

**Proof.** As shown in Fig. 2, there are a total of
ways to choose \( n_1 + f_1 \) busy or faulty input-middle interstage links and \( n_2 + f_2 \) busy or faulty middle-output interstage links. We can construct \( k \) pairs of overlapped interstage links in the following way: Initially, we select the \( n_1 + f_1 \) input-middle interstage links from a total of \( m \) input-middle interstage links and there are

\[
\binom{m}{n_1 + f_1}
\]

ways to do this; then, \( k \) busy or faulty input-middle interstage links which are overlapped with \( k \) busy or faulty middle-output interstage links can be chosen from the \( n_1 + f_1 \) input-middle interstage links and there are

\[
\binom{n_1 + f_1}{k}
\]

ways to do this; finally, we must select the rest of the \( n_2 + f_2 - k \) busy or faulty middle-output interstage links from the remaining \( m - n_1 - f_1 \) busy or faulty input-middle interstage links and there are

\[
\binom{m - n_1 - f_1}{n_2 + f_2 - k}
\]

ways to do this. Therefore, the probability that \( k \) pairs of links are overlapped is

\[
\frac{\binom{m}{n_1 + f_1} \binom{n_1 + f_1}{k} \binom{n_2 + f_2}{n_2 + f_2 - k}}{\binom{m}{n_1 + f_1} \binom{n_1 + f_1}{k} \binom{m - n_1 - f_1}{n_2 + f_2 - k}}.
\]

The probability can also be attained symmetrically by constructing the \( k \) busy or faulty input-middle interstage links which are overlapped with \( k \) busy or faulty middle-output interstage links by initially selecting \( n_2 + f_2 \) middle-output interstage links from a total of \( m \) middle-output interstage links, then the \( k \) overlapped links from \( n_2 + f_2 \), and, finally, the remaining \( n_1 + f_1 - k \) from \( m - n_2 - f_2 \).

We have determined the probability that \( k \) pairs of links are overlapped. We now establish the relationship between this fact and the probability that a connection request is not blocked. A connection request from input switch \( i \) to output switch \( j \) is not blocked if there exists at least one path from input switch \( i \) to output switch \( j \) in which both the input-middle and middle-output interstage links are not busy and are functional. This condition can be represented by

\[
n_1 + n_2 + f_1 + f_2 - k < m,
\]

which implies

\[
k \geq n_1 + n_2 + f_1 + f_2 - m + 1
\]

and

\[
k \geq \max\{0, n_1 + n_2 + f_1 + f_2 - m + 1\}.
\]

Having established a lower bound for the \( k \) overlapped links, it is obvious from Fig. 2 that

\[
k \leq \min\{n_1 + f_1, n_2 + f_2\}.
\]

Therefore, the probability that a connection is not blocked, given events \( n_1, f_1, n_2, f_2 \), and \( f_2 \) is given by

\[
\Pr\{\text{connection not blocked} \mid n_1, n_2, f_1, f_2\} = \min\{n_1 + f_1, n_2, f_2\} \left(\frac{1}{k+1}\right)^{m-n_2-f_2-k}.
\]

(6)

Given (6), we now need to determine the probability of the simultaneous occurrence of the four events \( n_1, f_1, n_2, f_2 \). Since there is dependency between events \( n_1 \) and \( f_1 \), and between events \( n_2 \) and \( f_2 \), we cannot simply assume that the four events are independent. However, we can group events \( n_1 \) and \( f_1 \) together, and events \( n_2 \) and \( f_2 \) together. Similar to previous work [13], [14], in order to make the calculation possible it is reasonable to assume that the occurrence of events \( n_1 \) and \( f_1 \) is independent of the occurrence of events \( n_2 \) and \( f_2 \) for sufficiently large \( n \) and \( r \).

Under this assumption we have

\[
\Pr\{n_1, n_2, f_1, f_2\} = \Pr\{n_1, f_1\} \cdot \Pr\{n_2, f_2\}.
\]

(7)

Let us calculate \( \Pr\{n_1, f_1\} \) and \( \Pr\{n_2, f_2\} \). We know that we can have at most \( m \) input-middle (or middle-output) interstage links which may be busy or faulty. Furthermore, it is evident that an interstage link cannot be both busy and faulty, which implies that a busy link is functional. Therefore, an interstage link can have three possible states:

1. faulty
2. functional and busy
3. functional and idle.

Let \( p_f \) denote the probability that an interstage link is faulty, \( p_b \) denote the probability that an interstage link is functional and busy, and \( q \) denote the probability that an interstage link is functional and idle. \( p_f \) may be specified by the network designer and can be determined by examining historical data on interstage link failures. However, as we shall see, a constraint exists which must be satisfied. Also, it is expected that \( p_f \) will be rather small, as larger values indicate an increasingly unreliable network. Furthermore, we need to ensure that \( p_b + p_f \leq 1 \) because an interstage link cannot be both busy and faulty. As in [18] and [14], let \( a \) be the probability that a typical input or output port is busy. We assume that the faults are uniformly distributed over \( m \) interstage links and the incoming traffic is uniformly distributed over all functional interstage links. Thus, for a given \( p_f \), the average number of functional interstage links is

\[
(1 - p_f)m = \frac{a}{(1 - p_f)m}.
\]

Clearly, \( p_b \) is a function of \( p_f \). We are interested in the small values of \( p_f \) and consider the case \( \frac{a}{(1 - p_f)m} \leq 1 \). Furthermore, it is important that values for \( n, m, a, \) and \( p_f \) conform to the constraint \( p_b + p_f \leq 1 \), which implies

\[
p_b = \min\left\{\frac{a}{(1 - p_f)m}, 1\right\}.
\]

(8)
\[
\frac{an}{(1 - n)^m} + p_f \leq 1.
\]

Therefore, \( p_f \) must be chosen such that

\[
p_f \leq 1 - \sqrt[1 - n]{m}.
\]

(9)

In addition, we represent the probability that an interstage link is idle by

\[
q = 1 - p_0 - p_f.
\]

(10)

To start with, it is reasonable to assume that the number of functional and faulty input-middle intersstage links (or middle-output intersstage links) follows a binomial distribution and the number of busy and idle input-middle intersstage links (or middle-output intersstage links) also follows a binomial distribution. However, given the fact that, among the \( m \) input-middle (or middle-output) intersstage links, at most \( n - 1 \) of them can be busy, we can obtain a more accurate probability distribution. Therefore, we represent the probability of the joint occurrence of events \( n_1 \) and \( f_1 \) as

\[
\Pr\{n_1, f_1\} = \sum_{j=0}^{n-1} \sum_{i=0}^{m-j} \binom{m}{j} \binom{n-j}{i} p_f^i p_0^{m-i-j} q^{n-1-j}.
\]

(11)

and the probability of the joint occurrence of events \( n_2 \) and \( f_2 \) as

\[
\Pr\{n_2, f_2\} = \sum_{j=0}^{n} \sum_{i=0}^{m-j} \binom{m}{j} \binom{n-j}{i} p_f^i p_0^{m-i-j} q^{n-j}.
\]

(12)

Given (6)-(12), we can now obtain the probability that a connection request is not blocked, accounting for intersstage link failures. It is given by

\[
P_{\text{bl}} = 1 - \{\Pr \text{ connection not blocked}\}.
\]

(13)

Note that in (13) the summation indices for \( f_1 \) and \( f_2 \) are from \( 0 \) to \( m - n_1 \) and from \( 0 \) to \( m - n_2 \), respectively. This is to account for the constraints \( 0 \leq n_1 + f_1, n_2 + f_2 \leq m \) and \( 0 \leq n_1, n_2 \leq n - 1 \).

Of particular interest in the above model is the special case \( p_f = 0 \), where the probability that an intersstage link is faulty is equal to 0, that is, a fault-free network. Let’s define \( \theta = 1 \). Then, when \( p_f = 0 \), (13) becomes

\[
P_{\text{bl}} = 1 - \{\Pr \text{ connection not blocked}\}.
\]

(15)

Thus, if \( p_f = 0 \), (13) reverts to the model for the non-fault-tolerant Clos network in (3).

5 DISCUSSIONS ON THE ANALYTICAL MODEL

In this section, we take a closer look at the new analytical model, under several network configurations. We are primarily interested in the effects of both an increasing failure rate and an increasing number of middle stage switches, given a constant failure rate, on the blocking probability given by (13).

Fig. 5 plots the blocking probability for three network configurations: \( n = r = 16 \) with \( 16 \leq m \leq 28 \), \( n = r = 32 \) with \( 32 \leq m \leq 48 \), and \( n = r = 64 \) with \( 64 \leq m \leq 84 \) at a network utilization of 80 percent for four different intersstage link failure rates. From Fig. 5, we observe that, for all network configurations, given a constant number of middle stage switches and a constant network utilization, as the probability of intersstage link failure increases, the blocking probability increases. For smaller intersstage link failure rates, say, \( p_f \leq 0.03 \), the blocking probability increases only slightly compared with that of \( p_f = 0 \). This indicates that the network has a good fault-tolerant capability in this case. However, the increase in blocking probability is more dramatic for larger values of \( p_f \) (\( > 0.03 \)).

Furthermore, in Fig. 5, we show the effect of an increasing number of middle stage switches (\( m \)) on the blocking probability given by (13). For the four curves (\( p_f = 0.0, p_f = 0.01, p_f = 0.03 \), and \( p_f = 0.07 \)) shown for each plot of Fig. 5, we see that as the number of middle stage switches increases, the blocking probability decreases sharply. Each plot also indicates that for two probabilities of intersstage link failure, \( p_f_1 \) and \( p_f_2 \) if \( p_f_1 > p_f_2 \) then equivalent blocking probability occurs when the number of middle stage switches for the curve for \( p_f_1 \) is greater than the number of middle stage switches for the curve for \( p_f_2 \). For all network configurations, given a constant intersstage link failure probability and a constant network utilization, increasing the number of middle stage switches results in decreasing blocking probability.

Finally, in Table 1, we give more data for different network configurations and different values of \( p_f \), which also shows the same trends as discussed for Fig. 5.

6 EXPERIMENTAL SIMULATIONS

An experimental study was performed to verify the approximations for the blocking probabilities given by the new analytical model (13). We developed a network simulator, which employed a random routing algorithm and allowed for intersstage link failure, to determine the blocking probability of a Clos network given the same set of network conditions used to derive the analytical model.

6.1 Assumptions

The network for which simulation data was generated is a three-stage Clos network. Comparisons between analytical data and simulated data were based on similar network configurations. A network configuration is defined to be a unique set of five variables: \( r \), the number of input stage and output stage switches in the network, \( n \), the number of input/output ports on each input/output switch, \( m \), the number of middle stage switches in the network, \( a \), the network utilization, and \( p_f \), the probability that an intersstage link in the network is faulty.

6.2 The Network Simulator

The network simulator consists of two main parts: a request generator and a request processor. The request generator
\[ \Pr \{ \text{connection not blocked} \} = \sum_{n_1=0}^{n-1} \sum_{n_2=0}^{n-1} \min(n_1, n_2) \sum_{m_1=0}^{n_1} \sum_{m_2=0}^{n_2} \frac{\binom{n_1}{k} \binom{m-n_1-k}{j} \binom{m}{j} \binom{m-f_1}{f_1} \binom{m-f_2}{f_2} \binom{m}{n_1} \binom{m}{n_2}}{\binom{m}{n_1+n_2+f_1+f_2-m-1} \binom{m}{n_1+n_2+f_1+f_2-m-1}} p_0^j p_1^j p_2^j q^{m-j} \]

(13)

Fig. 3. Equation (13).

\[ \Pr \{ \text{connection not blocked} \} = \sum_{n_1=0}^{n-1} \sum_{n_2=0}^{n-1} \min(n_1, n_2) \sum_{m_1=0}^{n_1} \sum_{m_2=0}^{n_2} \frac{\binom{n_1}{k} \binom{m-n_1-k}{j} \binom{m}{j} \binom{m-f_1}{f_1} \binom{m-f_2}{f_2} \binom{m}{n_1} \binom{m}{n_2}}{\binom{m}{n_1+n_2+f_1+f_2-m-1} \binom{m}{n_1+n_2+f_1+f_2-m-1}} p_0^j p_1^j p_2^j q^{m-j} \]

(15)

Fig. 4. Equation (15).

randomly generates a list of connection requests based on \( r \), the number of input and output stage switches, and \( u \), the number of input and output ports. A connection request is a four-tuple specifying the input port, the input-stage switch, the output-stage switch, and the output port. For routing purposes, we need only to be concerned with the input and output stage switches of the connection request. However, because we need to determine the legality of a connection request, the request generator must additionally generate the input and output ports of the connection request. The list of connection requests is generated before the actual simulation commences, as it is used as input to the request processor, which processes the list of connection requests. To strengthen comparisons among the simulation results, for a given \( n \) (and, likewise, \( r \)), the request processor utilized the same list of connection requests generated by the request generator.

Upon simulation startup, to account for interstage link failures, a certain portion of the interstage links are marked as faulty. This is accomplished for each interstage link by randomly generating a real number \( x \) such that \( 0 \leq x \leq 1 \). If \( x \leq p_f \), the interstage link is marked as faulty and cannot be used to route connection requests. These failures are considered permanent. To maintain a constant network utilization, as specified by \( u \), the network simulator must release active connections from the network. This is accomplished by the request processor when processing a legal and satisfiable connection request. When the network utilization is at the prescribed utilization, after establishing a connection request in the network, the request processor randomly chooses an active connection for termination. By doing so, the network utilization is held constant throughout the rest of the simulation. Finally, the blocking probability for a network simulation is defined to be the number of blocked connection requests divided by the total number of legal connection requests generated.

6.3 Methodology and Network Configurations

We were primarily interested in analyzing the approximations for the blocking probability of the analytical model to
see the effect of interstage link failure, using zero interstage link failures \((p_f = 0)\) as a basis. We used data generated by simulation to verify that the analytical model is consistent with repeated trials.

We examined three network configurations: \(n = r = 16\) with \(16 \leq m \leq 45\), \(n = r = 32\) with \(32 \leq m \leq 80\), and \(n = r = 64\) with \(64 \leq m \leq 140\). For each network configuration, we let \(50\% \leq a \leq 80\%\) in increments of 10 percent and we let \(p_f = \{0, 0.001, 0.005, 0.01, 0.03, 0.07\}\). For the simulations, \(a\) represents the network utilization and \(p_f\) represents the probability that an interstage link is faulty. Therefore, blocking probabilities were obtained for the same percentages. Furthermore, the request processor processed 20,000 legal connection requests from a list of 2,000,000 connection requests generated by the request generator.

We chose values for \(m\) which would allow comparison to the Clos deterministic nonblocking condition, \(m \geq 2n - 1\). Values for \(p_f\) were chosen on the basis of preliminary experimentation which suggested larger values of \(p_f\) yielded blocking probabilities indicative of a network with a minimal connecting capability. Of particular interest are the approximations for blocking probability given \(0 \leq p_f \leq 0.01\), as these values represented failure rates which would seem to be the most practical.

### 6.4 Simulation Results

In this section, we present the data generated for the Clos network simulations utilizing a random routing algorithm. We were primarily interested in verifying that the two main hypotheses discussed in Section 5 were consistent with repeated trials.

Table 2 shows the blocking probabilities resulting from the experimental simulations for similar network configurations presented for the analytical model in Table 2. Also, the blocking probability curves in Fig. 6 show the relationship between the analytical model data and the data resulting from the network simulations. First, for all values of \(n\) and \(r\), we see that the curves for the analytical model show the same trend as the curves for the network simulations, indicating that the analytical model is a good approximation of the actual blocking probability obtained by repeated trials. More specifically, those simulation results confirm our two main hypotheses discussed in Section 5 for the analytical model.
TABLE 1
Blocking Probabilities from Model Data for $n = 8$, $n = r = 16$, $n = r = 32$, and $n = r = 64$

<table>
<thead>
<tr>
<th>$p_f$</th>
<th>$m = 16$</th>
<th>$m = 20$</th>
<th>$m = 24$</th>
<th>$m = 28$</th>
<th>$m = 32$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>$4.922 	imes 10^{-1}$</td>
<td>$2.281 	imes 10^{-2}$</td>
<td>$1.964 	imes 10^{-4}$</td>
<td>$6.955 	imes 10^{-6}$</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0001</td>
<td>$4.923 	imes 10^{-1}$</td>
<td>$2.282 	imes 10^{-2}$</td>
<td>$1.961 	imes 10^{-4}$</td>
<td>$6.958 	imes 10^{-6}$</td>
<td>$2.339 	imes 10^{-14}$</td>
</tr>
<tr>
<td>0.005</td>
<td>$5.232 	imes 10^{-1}$</td>
<td>$3.349 	imes 10^{-2}$</td>
<td>$2.816 	imes 10^{-4}$</td>
<td>$7.862 	imes 10^{-7}$</td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>$5.547 	imes 10^{-1}$</td>
<td>$3.583 	imes 10^{-2}$</td>
<td>$4.060 	imes 10^{-4}$</td>
<td>$5.247 	imes 10^{-10}$</td>
<td></td>
</tr>
<tr>
<td>0.03</td>
<td>$6.825 	imes 10^{-1}$</td>
<td>$7.199 	imes 10^{-2}$</td>
<td>$1.451 	imes 10^{-3}$</td>
<td>$3.986 	imes 10^{-8}$</td>
<td></td>
</tr>
<tr>
<td>0.07</td>
<td>$9.993 	imes 10^{-1}$</td>
<td>$1.956 	imes 10^{-1}$</td>
<td>$1.956 	imes 10^{-2}$</td>
<td>$2.581 	imes 10^{-5}$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$p_f$</th>
<th>$m = 32$</th>
<th>$m = 40$</th>
<th>$m = 48$</th>
<th>$m = 56$</th>
<th>$m = 63$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>$2.697 	imes 10^{-4}$</td>
<td>$2.286 	imes 10^{-3}$</td>
<td>$4.479 	imes 10^{-7}$</td>
<td>$4.350 	imes 10^{-13}$</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0001</td>
<td>$2.704 	imes 10^{-4}$</td>
<td>$2.300 	imes 10^{-3}$</td>
<td>$4.540 	imes 10^{-7}$</td>
<td>$4.346 	imes 10^{-13}$</td>
<td>$6.140 	imes 10^{-14}$</td>
</tr>
<tr>
<td>0.005</td>
<td>$3.032 	imes 10^{-1}$</td>
<td>$3.057 	imes 10^{-3}$</td>
<td>$8.454 	imes 10^{-7}$</td>
<td>$2.181 	imes 10^{-12}$</td>
<td>$3.830 	imes 10^{-14}$</td>
</tr>
<tr>
<td>0.01</td>
<td>$3.392 	imes 10^{-1}$</td>
<td>$4.051 	imes 10^{-3}$</td>
<td>$1.540 	imes 10^{-6}$</td>
<td>$9.649 	imes 10^{-12}$</td>
<td>$6.051 	imes 10^{-14}$</td>
</tr>
<tr>
<td>0.03</td>
<td>$5.033 	imes 10^{-1}$</td>
<td>$1.150 	imes 10^{-2}$</td>
<td>$1.288 	imes 10^{-5}$</td>
<td>$6.883 	imes 10^{-11}$</td>
<td>$1.227 	imes 10^{-13}$</td>
</tr>
<tr>
<td>0.07</td>
<td>$8.539 	imes 10^{-1}$</td>
<td>$6.500 	imes 10^{-1}$</td>
<td>$3.804 	imes 10^{-1}$</td>
<td>$2.006 	imes 10^{-7}$</td>
<td>$9.786 	imes 10^{-11}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$p_f$</th>
<th>$m = 64$</th>
<th>$m = 84$</th>
<th>$m = 98$</th>
<th>$m = 115$</th>
<th>$m = 127$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>$7.334 	imes 10^{-2}$</td>
<td>$5.570 	imes 10^{-4}$</td>
<td>$2.240 	imes 10^{-6}$</td>
<td>$1.110 	imes 10^{-16}$</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0001</td>
<td>$7.369 	imes 10^{-2}$</td>
<td>$5.635 	imes 10^{-4}$</td>
<td>$2.585 	imes 10^{-6}$</td>
<td>$3.020 	imes 10^{-14}$</td>
<td>$3.086 	imes 10^{-14}$</td>
</tr>
<tr>
<td>0.005</td>
<td>$9.374 	imes 10^{-2}$</td>
<td>$9.968 	imes 10^{-4}$</td>
<td>$9.015 	imes 10^{-14}$</td>
<td>$1.585 	imes 10^{-13}$</td>
<td>$1.617 	imes 10^{-13}$</td>
</tr>
<tr>
<td>0.01</td>
<td>$1.160 	imes 10^{-1}$</td>
<td>$1.720 	imes 10^{-5}$</td>
<td>$2.694 	imes 10^{-14}$</td>
<td>$2.152 	imes 10^{-15}$</td>
<td>$2.121 	imes 10^{-13}$</td>
</tr>
<tr>
<td>0.03</td>
<td>$2.551 	imes 10^{-1}$</td>
<td>$1.360 	imes 10^{-4}$</td>
<td>$1.637 	imes 10^{-10}$</td>
<td>$4.369 	imes 10^{-15}$</td>
<td>$5.362 	imes 10^{-15}$</td>
</tr>
<tr>
<td>0.07</td>
<td>$7.316 	imes 10^{-1}$</td>
<td>$4.271 	imes 10^{-3}$</td>
<td>$1.298 	imes 10^{-7}$</td>
<td>$8.196 	imes 10^{-13}$</td>
<td>$8.562 	imes 10^{-13}$</td>
</tr>
</tbody>
</table>

Also, Fig. 6 confirms that for a constant network utilization and a constant number of middle stage switches, as the probability of link failure increases, the probability that a connection request cannot be satisfied increases. For small values of $p_f$, the increase in blocking probability is only within a narrow range. However, for larger values of $p_f$, the effect of intersstage link failure probability is felt more dramatically. Furthermore, Fig. 6 demonstrates that as $r$ (and likewise $n$) increases, the effect of intersstage link failure is increasingly evident for the network simulations. For $n = r = 16$ and $n = r = 32$ the curves for the network simulations are almost identical for $p_f = 0.005$ and $p_f = 0.01$. However, for $n = r = 64$ (and somewhat for $n = r = 32$) the blocking probabilities obtained for the smaller of the two link failure probabilities ($p_f = 0.005$) are lower than the blocking probabilities shown for $p_f = 0.01$ and we begin to see the two simulation curves look comparatively close to the two curves shown for the analytical model. This further demonstrates the consistency of the analytical model with repeated trials.

From Fig. 6, we also observe that there is some gap between the analytical results and the simulation results. This is mainly because that in the analytical model, every input stage switch is assumed to have exactly $(1-p_f)m$ functional intersstage links; but in the simulation, it is impossible to enforce this assumption. In fact, the number of functional intersstage links for each input stage switch may vary depending on where the faults are located. Thus, the simulation results which were obtained by averaging the blocking probabilities of all requests from different input stage switches are not exactly the same as those obtained by the analytical model. Since the intersstage link busy probability is inversely proportional to the number of functional intersstage links as shown in (8), the gap is more noticeable in the case of a larger $p_f$.

7 CONCLUSIONS
In this paper, we have presented a new analytical model for the routing blocking probability of the three-stage Clos network in the presence of intersstage link faults. Because the Clos network has a powerful connecting capability and is inherently more fault-tolerant of intersstage link failure than many proposed network designs, it continues to be an
### Table 2

Blocking Probabilities from Network Simulations for $n = r = 16$, $n = r = 32$, and $n = r = 64$ with $a = 0.8$ and $p_f = \{0, 0.001, 0.005, 0.01, 0.03, 0.07\}$

<table>
<thead>
<tr>
<th>$p_f$</th>
<th>$m = 16$</th>
<th>$m = 18$</th>
<th>$m = 20$</th>
<th>$m = 22$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>$3.266 \times 10^{-1}$</td>
<td>$7.820 \times 10^{-2}$</td>
<td>$9.900 \times 10^{-3}$</td>
<td>$5.000 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.001</td>
<td>$3.266 \times 10^{-1}$</td>
<td>$8.290 \times 10^{-2}$</td>
<td>$1.230 \times 10^{-2}$</td>
<td>$8.000 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.005</td>
<td>$3.680 \times 10^{-1}$</td>
<td>$9.610 \times 10^{-2}$</td>
<td>$1.330 \times 10^{-2}$</td>
<td>$7.000 \times 10^{-3}$</td>
</tr>
<tr>
<td>0.01</td>
<td>$3.790 \times 10^{-1}$</td>
<td>$8.420 \times 10^{-2}$</td>
<td>$1.740 \times 10^{-2}$</td>
<td>$1.500 \times 10^{-3}$</td>
</tr>
<tr>
<td>0.03</td>
<td>$4.414 \times 10^{-1}$</td>
<td>$4.408 \times 10^{-1}$</td>
<td>$2.840 \times 10^{-2}$</td>
<td>$4.000 \times 10^{-3}$</td>
</tr>
<tr>
<td>0.07</td>
<td>$7.655 \times 10^{-1}$</td>
<td>$2.925 \times 10^{-1}$</td>
<td>$8.580 \times 10^{-2}$</td>
<td>$2.480 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$p_f$</th>
<th>$m = 32$</th>
<th>$m = 36$</th>
<th>$m = 38$</th>
<th>$m = 40$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>$1.678 \times 10^{-1}$</td>
<td>$1.470 \times 10^{-2}$</td>
<td>$2.200 \times 10^{-3}$</td>
<td>$6.000 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.001</td>
<td>$1.687 \times 10^{-1}$</td>
<td>$1.500 \times 10^{-2}$</td>
<td>$2.600 \times 10^{-3}$</td>
<td>$3.500 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.005</td>
<td>$1.810 \times 10^{-1}$</td>
<td>$2.120 \times 10^{-2}$</td>
<td>$4.100 \times 10^{-3}$</td>
<td>$7.000 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.01</td>
<td>$2.182 \times 10^{-1}$</td>
<td>$2.310 \times 10^{-2}$</td>
<td>$5.000 \times 10^{-3}$</td>
<td>$6.500 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.03</td>
<td>$2.334 \times 10^{-1}$</td>
<td>$4.500 \times 10^{-2}$</td>
<td>$7.600 \times 10^{-3}$</td>
<td>$1.600 \times 10^{-3}$</td>
</tr>
<tr>
<td>0.07</td>
<td>$4.807 \times 10^{-1}$</td>
<td>$1.550 \times 10^{-1}$</td>
<td>$5.100 \times 10^{-2}$</td>
<td>$1.790 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$p_f$</th>
<th>$m = 64$</th>
<th>$m = 66$</th>
<th>$m = 68$</th>
<th>$m = 72$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>$4.070 \times 10^{-2}$</td>
<td>$1.480 \times 10^{-2}$</td>
<td>$5.550 \times 10^{-3}$</td>
<td>$2.500 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.001</td>
<td>$3.860 \times 10^{-2}$</td>
<td>$1.650 \times 10^{-2}$</td>
<td>$5.460 \times 10^{-3}$</td>
<td>$6.000 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.005</td>
<td>$4.870 \times 10^{-2}$</td>
<td>$1.910 \times 10^{-2}$</td>
<td>$5.850 \times 10^{-3}$</td>
<td>$7.000 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.01</td>
<td>$5.640 \times 10^{-2}$</td>
<td>$2.200 \times 10^{-2}$</td>
<td>$7.550 \times 10^{-3}$</td>
<td>$1.600 \times 10^{-3}$</td>
</tr>
<tr>
<td>0.03</td>
<td>$9.130 \times 10^{-2}$</td>
<td>$5.350 \times 10^{-2}$</td>
<td>$2.300 \times 10^{-2}$</td>
<td>$3.350 \times 10^{-3}$</td>
</tr>
<tr>
<td>0.07</td>
<td>$2.418 \times 10^{-1}$</td>
<td>$1.420 \times 10^{-1}$</td>
<td>$9.230 \times 10^{-2}$</td>
<td>$2.120 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

---

**ACKNOWLEDGMENTS**

Research was supported by the U.S. Army Research Office under Grant No. DAAH04-96-1-0234 and the U.S. National Science Foundation under Grant No. OSR-9350540.

---

**REFERENCES**


Fig. 6. Blocking probability of the Clos network comparing the analytical model results with the simulated results for $n = r = 16$ with $16 \leq n \leq 26$, $n = r = 32$ with $32 \leq n \leq 48$, and $n = r = 64$ with $64 \leq n \leq 88$ at 80 percent network utilization.


204-216, 1974.
Switched Networks Employing Originating Office Control with
June 1978.
[17] Y. Min, H.Y. Youn, "Performance Modeling and Evaluation of
Circuit Switching Using Clos Networks," IEEE Trans. Computers,
1998.
[21] Y. Yang and J. Wang, "Wide-Sense Nonblocking Clos Networks
[23] Y. Yang and C.M. Masson, "Nonblocking Broadcast Switching
Networks," IEEE Trans. Computers, vol. 40, no. 9, pp. 1,005-1,015,
Three-Stage Clos Networks," Networks, vol. 23, pp. 427-439, John 

Matthew P. Haynos received the BA degree in
computer sciences/applied mathematics and
cognitive science (honors) from the University of
Rochester in 1990, and the MS degree in
computer science from the University of Ver-
mont in 1998. He has been employed by the IBM
Corporation since June 1990. His research
interests include interconnection networks,
fault-tolerant computing, algorithm design, and
distributed computing.

Yuanyuan Yang received the BEng and MS
degrees in computer engineering from Tsinghua
University, Beijing, China, in 1982 and 1984,
respectively, and the MSE and PhD degrees in
computer science from Johns Hopkins Univer-
sity, Baltimore, Maryland, in 1989 and 1992,
respectively. She is currently an associate profes-
sor of computer engineering at the State
University of New York at Stony Brook. Before
joining SUNY Stony Brook, Dr. Yang was a
faculty member in the Department of Computer Science, University of
Vermont, in Burlington, from 1992-1999 (as an associate professor from
1998-1999). Dr. Yang's research interests include parallel and dis-
tributed computing and systems, high speed networks, optical networks,
high performance computer architecture, and fault-tolerant computing.
She has published extensively in major journals and refereed conference
proceedings related to these research areas. Dr. Yang holds two U.S. patents in the area of multicast communication networks,
with four more patents pending. Her research has been supported by the
U.S. Army Research Office and the U.S. National Science Foundation.
She has served on the program/organizing committees of a number of
international conferences. Dr. Yang is a senior member of the IEEE and
a member of the ACM, IEEE Computer Society, and IEEE Commu-
nication Society.